Computational Competitive and Negative Design to Derive a Specific cJun Antagonist

Alexander Lathbridge[1] and Jody M. Mason[1,2]

[1]Department of Biology & Biochemistry, University of Bath, Claverton Down, Bath, BA2 7AY, UK

[2]Corresponding author: Jody M. Mason (E: j.mason@bath.ac.uk; T: +441225386867)

Keywords: Coiled coil; Activator Protein-1; transcription factor; computational design, peptide libraries, protein-protein interactions

Basic leucine-zipper (bZIP) proteins reside at the end of cell-signalling cascades and function to modulate transcription of specific gene targets. bZIPs are recognised as important regulators of cellular processes that include cell growth, apoptosis and cell differentiation. One such validated transcriptional regulator, Activator Protein-1 is typically comprised of heterodimers of Jun and Fos family members, and is key in the progression and development of a number of different diseases. The best described component, cJun is upregulated in a variety of diseases that include cancer and osteoporosis and psoriasis. Towards our goal of inhibiting bZIP proteins implicated in disease pathways, we here describe the first use of a novel *in silico* peptide-library screening platform that facilitates the derivation of sequences exhibiting high affinity with cJun while disfavouring homodimer formation or formation of heterodimers with other closely related Fos sequences. In particular, using Fos as a template, we have computationally screened a peptide library of over 60 million members and ranked hypothetical on/off target complexes according to predicted stability. This resulted in the identification of a sequence that bound cJun, but displayed little homomeric stability or preference for cFos. The computationally selected sequence maintains similar interaction stability to a previous experimentally derived cJun antagonist, while providing much improved specificity. Our study provides new insight into the use of tandem *in silico* screening / *in vitro* validation and the ability to create a peptide that is capable of satisfying conflicting design requirements.

INTRODUCTION

Coiled coils (CCs) are located within 3-5% of all amino acid structures and are highly versatile in the interactions they drive. They are characterised by a repeat of seven amino acids, the heptad repeat, with a preference for particular residue types at each position[1]. Despite the apparent straightforward link from their primary sequence to quaternary structure, CCs are highly specific in driving a wide variety of diverse protein-protein interactions, making them highly relevant systems in biotechnology and synthetic biology, and as pharmacological targets. Parallel dimeric CCs structures found within bZIP (basic leucine zipper) motifs are one of the simplest examples – it is comprised of 2 left-handed supercoiled α-helices which intertwine via a large interacting surface area.

Efforts are on-going to be able to predict bZIP stability[2–5], and more recently specificity[2,4,6–8]. The ability to predict stability and specificity of peptides directly from the primary structure is a longstanding goal, and is particularly important given the large number of human bZIPs. Current attempts to design peptides that specifically inhibit target leucine zipper interactions have taken a incremental library-based approach, with each new attempt increasing our understanding of the factors that underpin overall affinity. For example, *in vitro* assays of the binding affinity of 53 human bZIPs[1] showed there to be multiple interaction profiles, with specificity both within and between discrete bZIP families. This amounts to over 1400 potential interactions and selectivity within this set of bZIP interactions demonstrates how inherent sequence elements govern the selectivity of CC interactions. We have previously utilised an intracellular library screening approach to derive specific antagonists of the oncogenic transcriptional regulator, Activator Protein-1 (AP-1).

Transcription factors represent compelling if difficult drug-targets by conventional small molecule approaches. Their modulation can ensure that erroneous signals can be blocked at the transcriptional level, thus halting production of target genes implicated in disease, irrespective of the upstream signal imposed. Indeed, many oncogenic signal transduction cascades are known to upregulate transcription factor activity, leading to gene expression changes that drive cell transformation[9-10] and that place bZIP families centre stage as therapeutic targets in cancer. Our previous efforts in this area have resulted in antagonism of AP-1 components by designing Jun or Fos-based peptide libraries. This has been followed by their expression and screening inside living cells for an interaction with their target protein[3,11,12]. In addition, we have experimented with methodologies in which competing off-target proteins are expressed in the assay during library selection. Both our conventional intracellular *Protein-fragment Complementation Assay* (PCA) library screening approach[3,11] as well as a target-specificity enhancing *Competitive and Negative Design*

Initiative (CANDI)[13] have resulted in a many PPI inhibitors[14,15] and CC forming peptides in which the target is sequestered from binding to its natural partner. These assays have the significant advantage over *in vitro* approaches, of enriching for target-specific sequences (relative to a broad range of other proteins expressed within the cell), as well as sequences that are structured, soluble, non-toxic, and that resist protease breakdown. Moreover the significant amount of data gained from these experiments, and our consequent increased understanding of the system, has facilitated the development of a series of tools that can work by predicting the affinities, and consequently specificities of CCs, based only on input of the primary amino acid sequences.

The bZIP CC Prediction Algorithm (bCIPA)[3,16] works by analysing the helical propensities of component helices, together with the predicted contribution from electrostatic interactions and core residue interactions, to estimate the thermal denaturation temperature (T_m) of all hypothetical dimeric species within a defined system. Driven by coupling energies that describe $\mathbf{a_i}$-$\mathbf{a'_i}$ hydrophobic interactions and $\mathbf{g_i}$-$\mathbf{e'_{i+1}}$ electrostatic interactions, as pairwise interactions measured by previous double mutant analysis studies in CCs[17,18]. bCIPA was derived to estimate the T_m of a given parallel dimeric CC using only the primary sequence and was shown to correctly predict 97% of all strong interactions and 95% of all non-interacting pairs using an independent data set of human bZIP proteins. As with previous prediction models, this approach allows prediction in a pairwise manner.

From previous work conducted to benchmark the accuracy of various prediction approaches using binding data from a FRET assay of interacting bZIP proteins [19], bCIPA has similar accuracy to other purely data-driven models[2,18] and had more prediction accuracy for the 948 experimentally derived binding values than models which were driven solely by (or in combination with) CC structural prediction[20,21]. Indeed, using the bCIPA engine we have recently screened very large peptide interactomes to identify sets of up to sixteen *de novo* designed peptides that when combined are capable of forming specific CCs with their cognate partners[6,7].

Building on these previous findings, here we describe our efforts to take the approach much further by describing the first utilisation of an *in silico* approximation to the PCA and CANDI-PCA approach, which we term *in silico* PCA (isPCA) and *in silico* CANDI (isCAN). The first approach allows the user to ***i)*** define a target ***ii)*** define every library member as a potential homodimeric off-target (Fig 1). The isCAN approach brings the additional capability of ***iii)*** entering user-defined sequences that can interact with either the target or library member. Both isPCA and isCAN allows the user to create and *in silico* screen a much more expansive library than is accessible experimentally using either the complementary intracellular PCA or CANDI-PCA approaches (i.e. ~10^7 for isCAN *vs*. ~10^5 for PCA). The software then selects sequences on the basis of the greatest ΔT_m between all non-desired states and

the desired target interaction, to give the highest predicted specificity. Here we describe the first implementation of the isCAN approach and provide an experimental validation of its use by computationally screening over 60 million peptides to identify candidates that can bind to cJun specifically in the presence of cFos.

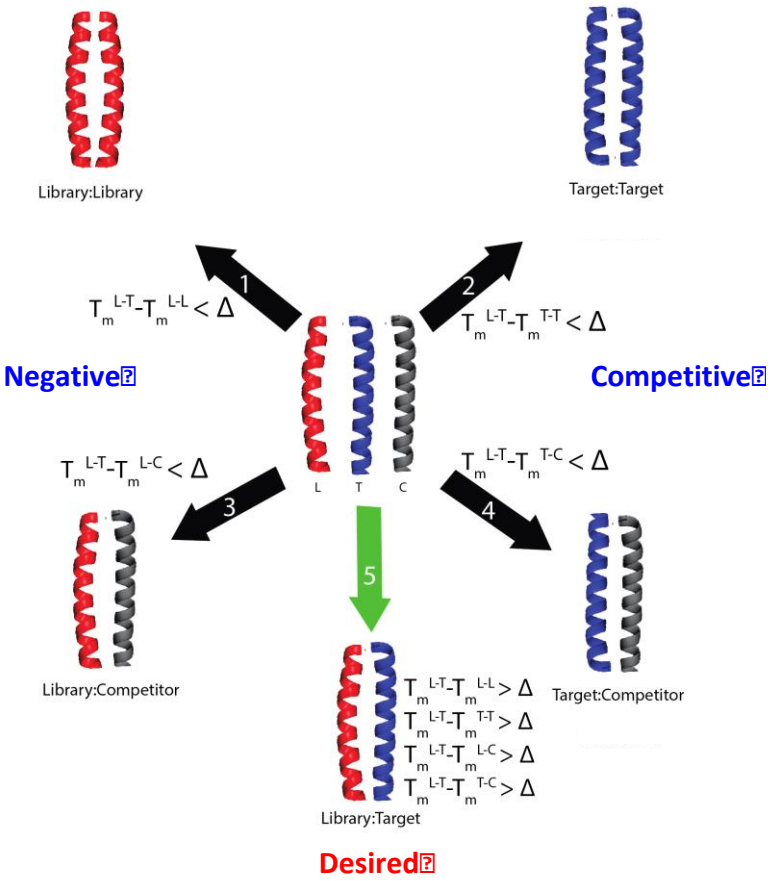

Figure 1 – Overview of the CANDI protocol. Shown are the desired and numerous undesired states that can form upon combination of the library/target/competitor peptides. Complexes 1, 2 (negative) and 5 (desired) are found within PCA, with the competitor complexes (4 and 5) introduced in CANDI. Within isCAN, specificity is driven by the desired delta value as specified by the user. The library member is only successful if it is able to form the desired complex with predicted T_m values greater than the delta.

Materials and Methods

In silico CANDI-PCA (isCAN) computationally screens a user-defined library against a given target. It identifies the highest predicted affinity binder to have greatest difference between target and off-target stability. This includes library homodimers as well as user-defined off-targets. isCAN utilises the underlying bCIPA algorithm[2-3], which incorporates Helical Propensity (HP), Core (C) and Electrostatic interactions (ES) to provide a quantitative estimate of the interaction affinities in the form of a thermal melting temperature (T_m) as follows:

$$Tm = (a \times HP) + (b \times C) + (c \times ES) + d \qquad (1)$$

The various functions within the algorithm assign scores to the peptide-peptide interactions. The size of the coefficients (a = 81.3256, b = -10.5716, c = -4.7771 and d= -29.1320) acts as a modifier for

the scale of the score. For the calculation of HP, the average α-helical propensity[4] of both peptides are calculated and totalled in Equation 2:

$$f(HP) = (\Sigma hp_a[l] + \Sigma hp_b[l]) \quad (2)$$

The nature of the frame alignment that bCIPA employs ensures that, if the peptides are not of same length, helical propensity is calculated to the length of the shorter peptide (l). For the calculation of the core interactions, only the residues within the hydrophobic interior (i.e. '***a***' or '***d***' *positions*) are considered such that the scoring mechanism is calculated accordingly:

$$f(core) = (\Sigma core_{res} \times y) \quad (3)$$

The format of this function is such that the non-considered residues are calculated but are disregarded (y = 0) . Only for the '***a***' and '***d***' positions is the value of y set to 1 (otherwise set to 0), ensuring that the core value in the final T_m calculation only incorporates these two heptad locations. As shown in Equation 4, only positions '***e***' and '***g***' are considered in an $\boldsymbol{g^i\text{-}e^{i+1}}$ parameter when calculating the electrostatic parameters:

$$f(ES) = (\Sigma es_{res}) \quad (4)$$

The program scans either peptide to calculate an electrostatic score for $\boldsymbol{g^i\text{-}e^{i+1}}$ and $\boldsymbol{e^i\text{-}g^{i-}1}$. This ensures that, even in the case of different length peptides, all of the potential electrostatic interactions are taken into account in the final ES score without unnecessarily incorporating a substituent score more than once.

Although the bCIPA engine has been previously employed to provide an *in silico* interactome prediction algorithm for the derivation of heterospecific coiled coil sets[6,7] all previous implementations of bCIPA have been restricted to estimating the T_m for single pairs of peptides forming a CC. Considering the ability that the algorithm displayed in accurately distinguishing interacting from non-interacting CCs, the logical next step was to expand its remit to mirror the semi-rational design and screening approach used in an experimental setting. isCAN simulates the CANDI extension of the PCA screening technique, and is a more advanced application of the bCIPA algorithm. The CANDI application of PCA involves the addition of competing peptides[13]. If the target or the library member favour complexes with the competitor peptide, cell growth is either reduced or halted (. 1). Similarly, isCAN is able to consider multiple off-targets in addition to the target. To achieve this, in addition to in-built frame alignment and prediction functions, isCAN has a number of unique in-built check points. These make use of the individual predictions relating to the library (L), target (T) and competitor (C) peptides. Due to the optimisation of core and electrostatic residues found in designed libraries, many peptides members are predicted to be more stable as

homodimeric complexes than as heterodimers with the target. isCAN is split into two sections – with the first set of calculations mirroring the PCA (isPCA section) and the second introducing the competitor peptides (isCAN). This stepwise calculation ensures that processing time is not wasted on library members which are predicted to preferentially homodimerise or are unable to overcome the target homodimer (and are therefore not "PCA successful"). A key concept in both is the predicted difference in T_m values (Δ). It is the key determinant behind the separation of successful and unsuccessful peptides in the library. User-defined, this value underlies all of the check-points that the software considers. In particular, if $\mathbf{T_m^{L\text{-}T} - T_m^{L\text{-}L} > \Delta}$ (i.e. the difference between the L-T desired heterodimer and the library homodimer is greater than the previously established desired delta value), then the peptide is considered homodimerically successful and proceeds to the subsequent stages. Any peptides that meet both this and the target homodimer delta ($\mathbf{T_m^{L\text{-}T} - T_m^{T\text{-}T} > \Delta}$) are considered "PCA-successful" (i.e. complying with the scenario found during a PCA). The PCA-succesful library members then have their desired state T_m ($\mathbf{T_m^{L\text{-}T}}$) compared with CANDI-specific competitive library off-target T_m values (i.e. $\mathbf{T_m^{L\text{-}T} - T_m^{T\text{-}C} > \Delta}$ and $\mathbf{T_m^{L\text{-}T} - T_m^{L\text{-}C} > \Delta}$) and the 'CANDI-successful' library members are next exported for further analysis. Due to the multiple ways in which AP-1 may actively form, users can enter other Fos and Jun family member sequences to impose target-specificity upon the screen. This addresses one of the key points of computationally aided peptide design with large families of peptides – avoiding interactions with other bZIPs which may be transcriptionally active and beneficial.

Calculating off-targets - In any simulated CANDI system, the interactions can be expressed as ***2n+3*** where *n* refers to the number of peptides introduced as competitive molecules, with only one desired (L-T) interaction. As an example, a cJun-targeting library would utilise four Fos family members (cFos, FosB, Fra1,Fra2) as competitors. As previously mentioned, this would result in 10 off-targets for each peptide. As such, this would be 10 T_m values that the single desired L-T complex T_m must be able to overcome with $\Delta > 0$. To maximise the efficiency of the tool, the initial calculations made by the prediction software are of the heterospecific (L-T) complex. The output of this simple screen is used to partition the library by desired T_m, irrespective of the ability to outcompete predicted off-target complexes. These partitions are 10^6 in size and are used to break down the computationally expensive isCAN calculations into computationally less demanding processes - approximately 36000 calculations are processed in one minute (with this value increasing over time as increasing amounts of calculation data is stored within the application).

Peptide Synthesis - Rink amide ChemMatrix™ resin was obtained from PCAS Biomatrix, Inc. (St.-Jean-sur-Richelieu, Canada); Fmoc L -amino acids and 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetra-methyluronium hexafluorophosphate or benzotriazol-1-yl-ox-ytripyrrolidinophosphonium

hexafluorophosphate were obtained from AGTC Bioproducts (Hessle, UK); all other reagents were of peptide synthesis grade and obtained from ThermoFisherScientific (Loughborough,UK).

Peptides were synthesised on a 0.1-mmol scale on a PCAS ChemMatrix™ Rink amide resin using a Liberty Blue™ microwave peptide synthesiser (CEM; Matthews, NC) employing Fmoc solid-phase techniques[22] with repeated steps of coupling, deprotection and washing (4 × 5 ml dimethylformamide).

Coupling was performed as follows: Fmoc amino acid (5 eq), 2-(1H-benzotriazole-1-yl)-1,1,3,3 tetramethyluronium hexafluorophosphate or benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate(4.5 eq) and diisopropylethylamine (10 eq) in dimethylformamide (5 ml) for 5 min with 35-W microwave irradiation at 90 °C.

Deprotection was performed as follows: 20% piperidine in dimethylformamide for 5 min with 30-W microwave irradiation at 80 °C. Following synthesis, we acetylated the peptide—acetic anhydride (3 eq) and diisopropylethylamine (4.5 eq) in dimethylformamide(2.63 ml) for 20 min—and then cleaved it from the resin with concomitant removal of side-chain-protecting groups by treatment with a cleavage mixture (10 ml) consisting of TFA (95%), triisopropylsilane (2.5%) and H_2O (2.5%) for 4 h at room temperature.

Suspended resin was removed by filtration, and the peptide was precipitated using three rounds of crashing in ice-cold diethyl ether, vortexing and centrifuging. The pellet was then dissolved in 1:1MeCN/H_2O and freeze-dried. Purification was performed by RP-HPLC using a Phenomenex Jupiter Proteo (C18) reverse-phase column (4 µm, 90 Å, 10 mm inner diameter × 250 mm long). Eluents used were as follows: 0.1% TFA in H_2O (a) and 0.1% TFA in ACN (b).

The peptide was eluted by applying a linear gradient (at 3.5 ml/min) of 5–95% B over 40 min. Fractions collected were examined by electrospray MS, and those found to contain exclusively the desired product were pooled and lyophilised. Analysis of the purified final product by RP-HPLC indicated a purity of >95%.

Circular Dichrosim (CD) was carried out using an Applied Photophysics Chirascan CD apparatus (Leatherhead, UK) using a 200-µl sample in a CD cell with a 1-mm path length. Samples contained 150 µM total peptide (Pt) concentration at equimolar concentration for heterodimeric solutions (i.e., 75 µM per peptide) and suspended in 10 mM potassium phosphate and 100 mM potassium fluoride at pH 7 for 30 minutes prior to analysis. The CD spectra of samples were scanned between 300 nm and 190 nm in 1 nm steps, averaging 0.5 s at each wavelength. Three scans at 20 °C were averaged

to assess helical levels and CC structure. Raw data (ellipticities) were collected and averaged, and data were converted to molar residue ellipticities (MRE).

Thermal denaturation experiments were performed at 150 µM in a buffer of 10 mM potassium phosphate and 100 mM potassium fluoride at pH 7. The instrument used was an Applied Photophysics Chirascan Circular Dichroism device. For all thermal denaturation experiments, a stepping gradient was set from 0°C to 90°C in 1°C increments (except for cFos containing complexes – which stopped at 50°C). Each temperature point was held for 0.5 min to equilibrate sample before scanning ellipticity at 222 nm. Melting profiles converted to equilibrium denaturation curves fitted to a two-state model, derived via modification of the Gibbs–Helmholtz equation to yield the melting temperature (T_m).[23]

Size-exclusion experiments chromatography experiments were performed at room temperature using a Superdex Peptide 10/300 GL column (GE Healthcare Life Sciences) by injecting 100 µl of a 150 µM Pt sample in 10 mM potassium phosphate and 100 mM potassium fluoride at pH 7 and at a flow rate of 0.5 ml/min. Elution profiles were recorded via A_{280}.

Results and Discussion

cFos based Library Generation. We previously utilised a number of *in vitro*[12,16] and *in cellulo*[3,11] peptide library screening approaches to derive sequences capable of binding to the cJun target of AP-1. One of these efforts utilised a PCA approach with a library of 62,000 members to result in a 37 residue cJun antagonist (FosW)[3]. Using FosW as a template for further library design, this was followed by a truncated variant, 4HFosW, that retained most of the interaction affinity[11]. More recently we have taken this further by rationally designing helix-constrained variants to permit downsizing of the molecule while retaining binding affinity[24]. As a mechanism for further increasing target selectively during selection we have also expressed off-target homologous sequences during PCA selection. The CANDI-PCA approach works by maximising the difference in free energy of binding between the target and off-target complexes by removing non-selective library members[13].

Here we present a powerful new approach, based on the bCIPA engine, to facilitate the *in silico* derivation of specific peptide antagonists. As an exemplar we have derived a 39-residue antagonist that is specific for the cJun target. The sequence incorporates two additional residues over earlier designs - one ***g*** one ***e*** residue located at the N-terminus and C-terminus respectively (Fig. 2A). These permit four additional electrostatic interactions and in doing so provide greater scope for stabilisation/destabilisation of target/off-target complexes to enhance interaction specificity. Using

this extended cFos sequence as our design scaffold (Fig. 2A), we have implemented an *in silico* approximation to the PCA and CANDI-PCA *in cellulo* screening systems to allow a rapid prediction of peptide sequences that display high target specificity. The tools described in this study are freely available (see supporting information).

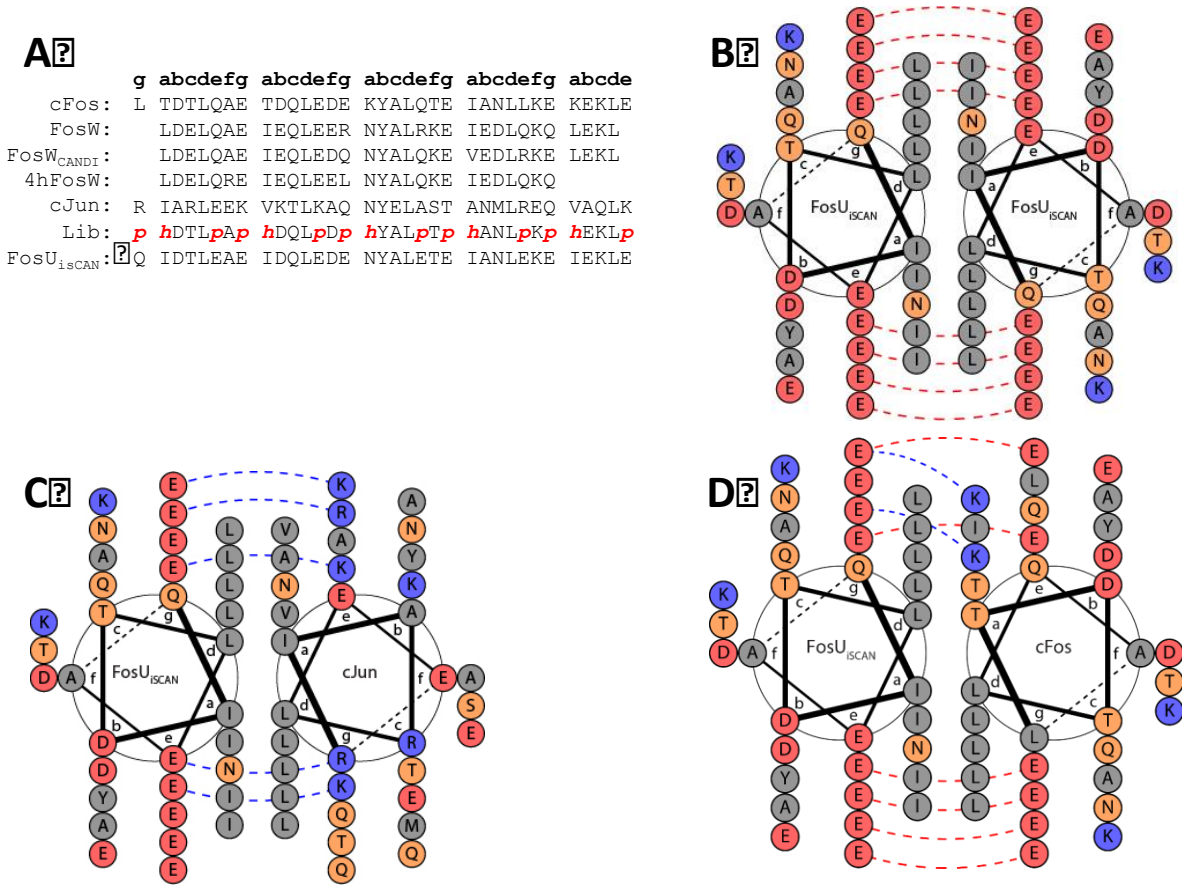

Figure 2 – Design of peptide inhibitor sequences (a). Peptide options are randomised around g/e positions (p) and a position (h). Compared to previous designs, the FosU$_{iSCAN}$ has had extensions added at N-terminal and C-terminal positions to add 2 extra residues for extra electrostatic interactions. The helical wheel diagrams, generated with DrawCoil 1.0[5], **(b-d)** display the residues present on the coiled coil from the position of the N-terminus to the C-terminus, looking down the axis of the alpha helices. These diagrams illustrate the hydrophobic interface at the core position (***a/d***) and the charged residues present at the flanking position (***e/g***). The repulsive residues found at the electrostatic positions in off-target complexes (b and d) are selected. The helical wheel of FosU$_{iSCAN}$ – cJun (c) demonstrates how FosU$_{iSCAN}$ has favourable electrostatic and core interactions to drive coiled coil formation.

During library generation, each position was inspected and options placed into the library sequence that corresponded to core hydrophobic and electrostatic positions within the heptad repeats (***a***, ***e*** and ***g***)**.** This resulted in an *in silico* library size of 60,466,176 peptides (**a** = ILVN, ***e*** = QEK, ***g*** = QEK, with N included at all ***a*** positions to give rise to potential specificity driving N-N pairs with the **a3** position on the target helix and to mitigate against the formation of higher-order oligomers [25–28]. The predicted ΔT_m parameter (defined as the difference between the T_m of desired dimer and closest non-desired dimer) was set to 20°C during parameter initialisation, since this was found to be the lowest value that resulted in a library able to be screened within a reasonable timeframe while retaining a large number of peptides predicted to be successful, such that library diversity was retained. We followed only members of the top 10^6 (partition 1) following these initial calculations, as this provided computationally efficiency (reducing calculation time from ~3 days to 5 hours - see supporting information) whilst selecting library members of the highest predicted target affinity. This step reduced the expansivity of the search prior to entry into the more stringent isCAN step, which additionally considered members of the cFos family that naturally interact with cJun (i.e. cFos, FosB, Fra1, Fra2) as explicit off-targets. As a competitive step in the initial isPCA, additional consideration of potential library members as homodimers and the stability of the cJun target as a homodimer were simulated (Fig. 1). During this step, many peptides formed predicted homodimers or were not suitably more stable than the predicted T_m of the homodimeric cJun target complex, and were consequently unable to overcome the stringent desired ΔT_m. Once the isPCA section was completed, the 60,466,176 member library was reduced to 73,124 peptides – a reduction to 0.12% of the original library. The predicted ΔT_m values for the PCA complexes drove this reduction, i.e. the difference between Library–cJun (L–T), and cJun–cJun (T–T) or Library–Library (L–L) interactions. Each successful peptide in the pool that satisfied the ΔT_m parameter set was permitted to proceed. These were described as sequences with predicted T_m values for non-desired states ($T_m^{L\text{-}L}$ and $T_m^{T\text{-}T}$) of at least 20°C less than the predicted Library-cJun T_m ($T_m^{L\text{-}T}$). These “isPCA-successful” peptides were next entered into isCAN. This final step introduced simulated competitors (in this case members of the Fos family that are known to form transcriptionally active complexes with cJun: cFos, FosB, Fra1 and Fra2). The isCAN step reduced the remaining library size further to 71,667 peptides. The isCAN step was critical in removing 1457 members that were predicted to bind to one of the Fos off-target competitor peptides ($T_m^{L\text{-}C}$). These were again defined as those unable to overcome the required ΔT_m values between the Library-cJun ($T_m^{L\text{-}T}$), the target – competitor complex (($T_m^{T\text{-}C}$) and Library –competitor ($T_m^{L\text{-}C}$) complexes. The average predicted ΔT_m for all L-C complexes (292,496 interactions) was 11°C.

Peptide Selection - FosU$_{isCAN}$. From the reduced size pool, peptides were finally ranked by the predicted ΔT_m value according to isCAN. This ensured that the peptide chosen for further study (FosU$_{isCAN}$) would exhibit both a high predicted T_m^{L-T} and a large ΔT_m (i.e. >20 °C) between this and the most stable of off-targets. As shown in Table 1, the peptide pool was reduced further to generate the top 10 sequences ranked by ΔT_m, which allowed for comparison of similarities and key differences between sequences.

isCAN#	Peptide *p* *h*DTL*p*A*p* *h*DQL*p*D*p* *h*YAL*p*T*p* *h*ANL*p*K*p* *h*EKL*p*	Library-Target T_m (°C)	Library-Library T_m (°C)	ΔT_m
1	Q IDTLEAE IDQLEDK NYALKTE LANLEKE IEKLE	92.5	39.3	52.0
2	K IDTLEAE IDQLEDK NYALKTE IANLEKE IEKLE	94.2	42.7	51.5
3	Q IDTLEAE IDQLEDK NYALKTE IANLEKE IEKLE	91.9	38.1	51.0
4	K IDTLEAE IDQLEDK NYALKTE LANLEKE IEKLE	94.8	43.9	50.9
5	K IDTLQAE IDQLEDK NYALKTE IANLEKE IEKLE	91.1	41.2	49.8
6*	Q IDTL**E**A**E** IDQL**E**D**E** NYAL**E**T**E** IANL**E**K**E** IEKL**E**	91.3	41.6	49.7
7	K IDTLQAE IDQLEDK NYALKTE LANLEKE IEKLE	91.8	42.4	49.3
8	Q IDTLEAE IDQLEDE NYALETE LANLEKE IEKLE	91.9	42.8	49.1
9	K IDTLEAE IDQLEDK NYALKTE NANLEKE IEKLE	88.2	30.6	48.0
10	K IDTLKAE IDQLEDK NYALKTE LANLEKE IEKLQ	88.4	35.9	48.0

Table 1. Top 10 peptides ranked by ΔT_m predictions calculated by isCAN screening, representing the top 0.01% of isCAN successful peptides . FosU$_{isCAN}$[6] was selected for validation (*and named FosU$_{ISCAN}$) due to the occurrence of Glu residues at ***g*** and ***e*** positions (in bold). These were predicted to have maximal beneficial interactions with cJun and maximal repulsion with off-targets (i.e. in complex with cFos and as a homodimer). Sequences additionally contain N-cap (AS) and C-cap (GAP) motifs not depicted here.

These peptides represented the top 0.01% of all peptides to successfully emerge from isCAN. The final sequence, FosU$_{isCAN}$6 (termed FosU$_{isCAN}$) was selected for validation with the rationale that the high level of similarity between the 10 sequences and corresponding T_m values and ΔT_m values from nearest stability off-targets. The selected sequence was chosen on the basis of 'charge blocks' – blocks of basic or acidic side chains at **e/g** positions that contribute favourably to the overall ΔT_m [6,7], but for which the sequence context of otherwise energetically equivalent residue contribution is not considered by the software (see also below). Additionally, the minimal difference between predicted

T_m and ΔT_m within sequences listed in Table 1 meant that the predictive power of the software should be able to be validated without using the top peptide. As shown in Figure 2B – 2D, this potential inhibitor was not expected to form interactions with off-targets and to be able to outcompete all possible other complexes (satisfying the competitive and negative design requirements of the experiment). As shown in Figure 3, the predicted ΔT_m between the closest off-target (in this case, L–L) and the desired (L–T) complex is 50°C. The predicted T_m for the $FosU_{isCAN}$-cJun interaction (91°C) is far greater than that of the closest undesired interactions (of 40°C for the cJun homodimer and 41°C for the $FosU_{isCAN}$ homodimer).

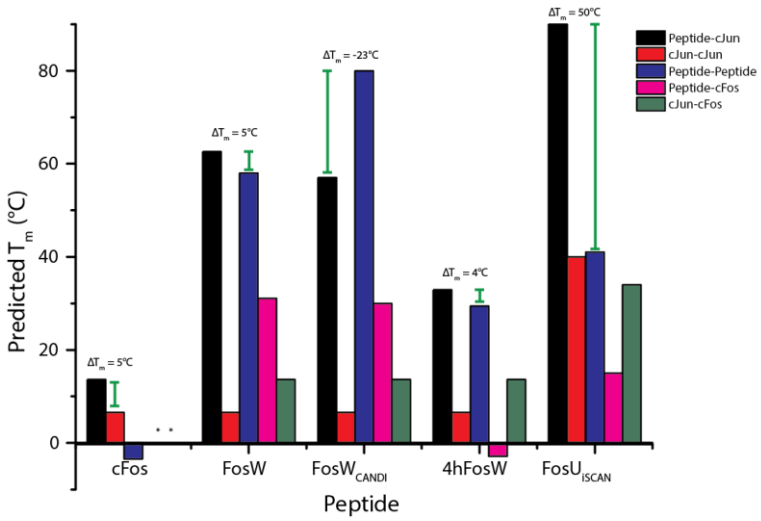

Figure 3 – Predicted T_m values of the isCAN selected peptide. $FosU_{isCAN}$ is compared against previous Fos-based peptides targeting cJun using a cFos competitor (and cFos, not duplicating values*). All interactions were predicted using the same isCAN protocol. The ΔT_m values against the highest off-target (predicted library member homodimerisation for all but cFos). $FosU_{isCAN}$ is predicted to have a T_m of 91°C with a ΔT_m of 50°C, both of these values are the highest predicted for Fos-based peptides targeting cJun.

isCAN Prediction. The isCAN selected sequence ($FosU_{isCAN}$: QIDTLEA EIDQLED ENYALET EIANLEK EIEKLE) was predicted to be structurally optimised for maximising and minimising desired and non-desired interactions, as shown using helical wheel diagrams (Fig 2B-2D). For the negative design in avoiding formation of the $FosU_{isCAN}$ homodimer complex, the electrostatic interactions play a vital role in destabilisation. This peptide resulted in the introduction of ***e/g*** charge blocks[6,7] which was previously shown to be important in driving intramolecular repulsion between neighbouring electrostatic side-chains. We previously found such charge patterns to further assist in concomitantly driving both favourable interactions between antagonist and target, as well as repulsions between potential antagonist homodimers (see also below), resulting in favourable gains in the measured ΔT_m. The introduction of these sequence-specific changes into antagonists otherwise considered energetically equivalent by the bCIPA approach can provide important contributions; they provide both intramolecular and intermolecular electrostatic contributions to

stability that can concomitantly stabilise the desired state while destabilising the homodimer. This is because neighbouring residues of same charge act to enhance or diminish the predicted energetics of intermolecular e/g interactions. This means that for homodimers the intramolecular repulsions act to enhance the intermolecular repulsions, making the complex less stable than is predicted without considering ***e/g*** residue sequence context. Concomitantly, for the desired heterodimer, the intermolecular repulsions act to assist the intermolecular attractions to a greater extent than predicted without ***e/g*** sequence context. The $FosU_{isCAN}$ electrostatic interactions provide intermolecular charge blocks of 4-5 residues at ***g/e*** positions, which serves to add additional destabilisation to $FosU_{isCAN}$ homodimer while adding additional stability to the target-bound heterodimeric complex. This is due to the presence of a Glu residue at both $\boldsymbol{g^2}$ and $\boldsymbol{e^3}$ – a combination not found in any other peptide within the top 10 from which $FosU_{isCAN}$ was selected. As shown in Table 1, many of those specific ***g*** and ***e*** positions were populated by Lys residues. For $FosU_{isCAN}$, of the 10 possible electrostatic interactions between the residues at positions ***e*** and ***g,*** 50% contain favourably charged profiles (i.e. negatively charged in $FosU_{isCAN}$ interacting with positively charged residues in cJun) and 50% have non-optimal profiles (negatively charged *vs.* neutral charge or hydrophobic). E.g. $\boldsymbol{e^2 – g'^1}$**,** $\boldsymbol{g^1 – e'^2}$ **and** $\boldsymbol{g^4 – e'^5}$ all contain electrostatically favourable salt bridge interactions between oppositely charged Glu – Lys. Similarly, $\boldsymbol{e^1 – g'^0}$ and $\boldsymbol{g^3 – e'^4}$ feature favourable Glu – Arg interactions. Non-optimal profiles are a result of selection against residues native to the ***e*** and ***g*** positions of cJun, where negatively charged Glu is facing Ala, Gln, Thr and Gln ($\boldsymbol{e'^3}$**,** $\boldsymbol{g'^2}$**,** $\boldsymbol{g'^3}$ **and** $\boldsymbol{g'^4}$**).** As shown in Table 1, inclusion of Gln at $\boldsymbol{g^0}$ for $FosU_{isCAN}$ is found in 3 other peptides within the top 10. As this position is facing a Glu at $\mathbf{e'^1}$ (or a Gln in all Fos family members considered as off-targets), 6 of the peptides within the top 10 favourably target this by selecting Lys. However, Lys would form favourable interactions within the homodimeric complex by forming a $\boldsymbol{g^0 – e'^1}$ Lys-Glu interaction (as in 6 other peptides found within the top 10). This is important because the cJun target peptide contains an Arg at $\boldsymbol{g'^0}$, meaning that the software has to decide on a $\boldsymbol{g^0}$ residue selection driven by the optimal interaction with this Arg while balancing potential off-target interactions and selecting the option that will overall contribute to the greatest ΔT_m. The bCIPA algorithm scores an Arg-Glu interaction more favourably (-2.0) than an Arg-Gln interaction (-1.5) or Arg-Lys interaction (-0.5). Therefore, it is more locally beneficial to have the $\mathbf{e^1}$ position filled by Glu. Thus, to avoid favourable interactions in the homodimeric complex, it is locally favoured to populate $\boldsymbol{g^0}$ with Gln rather than Lys/Glu. Although there is a high level of sequence similarity within it, there are no peptides within the top 10 which differ from $FosU_{isCAN}$ at a single residue switch at $\boldsymbol{g^0}$ to give full Glu at all ***e/g*** positions. What is observed instead is the introduction of Lys within different heptads at different peptides. This is due to the aforementioned non-optimal residues on native cJun

(***e'³, g'², g'³*** and ***g'⁴***) with Lys-Glu and Lys-Gln interactions being scored favourably (both contribute -1.5)[17,23]. The Lys-Ala interaction has no associated electrostatic contribution value and this is discussed below. This highlights part of the conflicting design requirements that isCAN attempts to address. At the core, cJun residues are optimised for hydrophobic interaction, with the ***a'*** position consisting of Ile, Val, Asn, Ala and Val. $FosU_{isCAN}$ takes advantage of this core arrangement with Ile at position ***a*** (with a3 as Asn to capitalise on the oligomer limiting locus of the ***a3*** N-N interaction[25]). Across the top 10 peptides, the major differences is ***a⁴***, where 40% of the peptides are Ile and 50% Leu (with 1 sequence selecting Asn) facing an Ala on ***a'⁴***. Both Ile/Leu contribute equally according to the algorithm (Ile-Ala /Leu-Ala = -0.5).

Formation of the complex with the competitor molecule cFos is predicted to be of low stability (Figure 2D), from both an electrostatic and a hydrophobic standpoint. cFos core is poorly optimised for hydrophobic interaction (compared to transcriptionally functional cJun) due to the presence of multiple polar and charged Thr/Lys residues. Similarly, 60% of the cFos-$FosU_{isCAN}$ electrostatic interactions are repulsive (Glu-Glu, +0.4 kcal/mol[17]). Moreover, the presence of Leu at **e4'** and **g0'** positions does not allow for further beneficial electrostatic interactions.

Circular Dichroism **-** An analysis of global secondary structure content for homodimeric and heterodimeric systems was conducted at 150 μM total peptide concentration to provide equimolar concentrations of each component helix for all dimeric systems. CD spectra showed $FosU_{isCAN}$ to exist as a 15.4% weakly populated helical structure (Fig 4) with the 208 nm signal significantly exceeding that of 222 nm. Similarly, cFos (Fig 4A) and cJun (Figure 4B) existed as 27.5% and 20.6% helical structures with 222/208 ratios of 0.60 and 0.82, respectively[29].

To analyse whether the selected peptide formed a complex with the cFos competitor sequence, a secondary structural scan of $FosU_{isCAN}$ – cFos was taken with CD (Fig 4A). As for component helices, this spectrum demonstrated the sample to lack both α-helical content (14.2%) and the double minima (222/208 ratio = 0.5), indicating that the two component helices are not able to associate into a CC. Interestingly, both monomers involved in this heteromeric system displayed greater α-helical content when measured in isolation.

In contrast, the secondary structure content of the $FosU_{isCAN}$ – cJun containing sample (Fig 4B) exhibited a much more intense signal with greater α-helical content (75.1%) – almost four times the signal of the constituent peptides in isolation. In addition the ratio of 222/208 was 0.98, providing further evidence for a significant increase to the helical stability of the sample. This demonstrates that the incubation of cJun with $FosU_{isCAN}$ elicits a significant conformational change in the sample and provides compelling evidence for the formation of a CC[30,31].

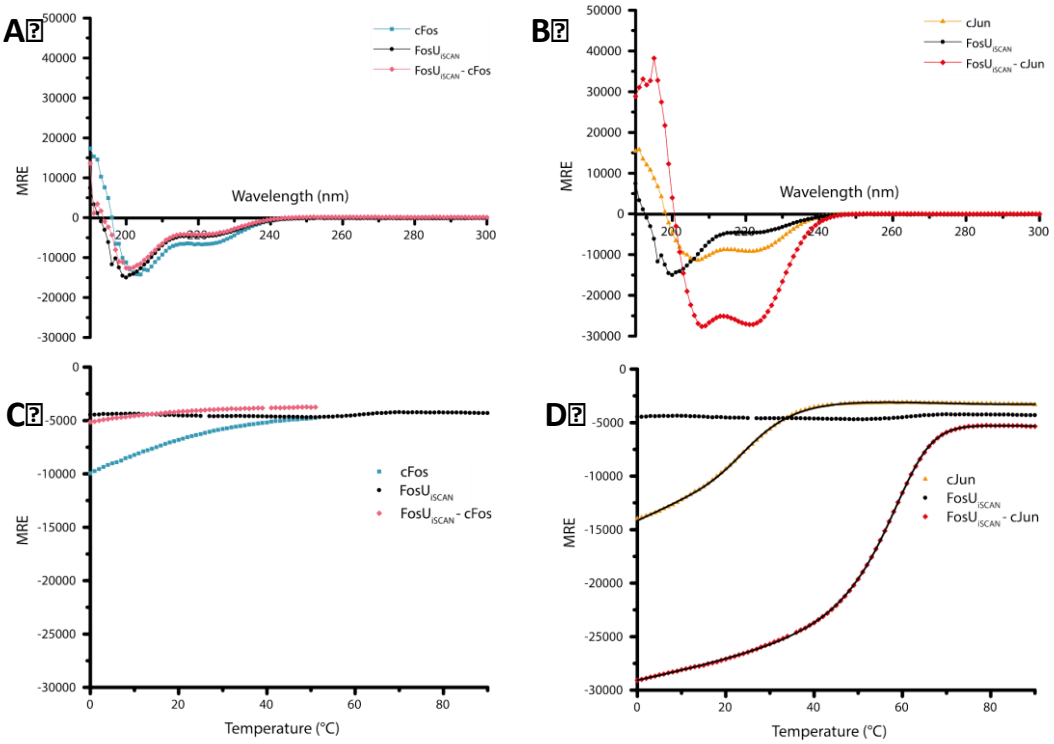

Figure 4 – **CD spectra and thermal denaturation data.** Shown are data for selected inhibitor peptide with cFos (a, c) and target cJun (b, d). Spectra were measured at 20°C at a total peptide concentration of 150 uM and presented as mean residue ellipticity (MRE). The minima at 208 nm and 222 nm are indicative of a helical structure, with the 222/208 ratio of the inhibitor with the cJun target showing more structure (222/208 = 0.98) over the undesired complex with cFos (222/208 = 0.56). This suggests that the inhibitor preferentially heterodimerises with cJun. Thermal denaturation profiles of homodimeric peptides and $FosU_{isCAN}$ with heterodimers (c/d) were taken using 1°C increments and tracking the 222nm signal at 150uM. $FosU_{isCAN}$ shows an increase in the transition mid-point when in complex with cJun (d), demonstrating a T_m of 57°C compared to the off-target state with cFos (c) which was unable to provide a measured T_m All experiments were undertaken in 10 mM potassium phosphate, 100 mM potassium fluoride at pH 7. Where possible (d), data was fitted to the two-state model.

Thermal Denaturation Profiles – Having observed a significant increase in the global secondary structure content of the cJun- $FosU_{isCAN}$ sample, we next sought to (Figure 4C/D and Table S2) quantify the stability of the complex by undertaking thermal denaturation experiments. In agreement with the spectra, this pattern of increased stability between undesired and desired complexes was also observed using thermal melts taken in 1°C increments. $FosU_{isCAN}$ in isolation did not form a CC complex – rather only the upper baseline characteristic of the profile was observed (Figure 4C - black). This is in agreement with spectral data and at is indicative of a weakly populated helix that does not self-associate. Further evidence for this is provided by size-exclusion chromatography (SEC) which demonstrates that the prominent species populated is monomeric (Figure 5A - blue). When $FosU_{isCAN}$ is incubated with with cFos (Figure 4C - pink), the thermal denaturation signal is similar to that of the component peptides. In contrast, for homomeric cJun (Figure 4D - orange), a clear transition mid-point is visible (27°C). Similarly, SEC experiments demonstrated cJun to exist as a dimer in solution (Figure 5B- red). However, when cJun was incubated with the $FosU_{isCAN}$ antagonist peptide, the helical signal increasesd significantly and led to an increased transition midpoint of 30 °C (Figure 4D - red)– demonstrating an increase in thermal stability to 57°C. This shift was further confirmed by SEC, demonstrating that the entire sample was in a dimeric state and that cJun had therefore paired with $FosU_{isCAN}$ (Figure 5B). The inability of $FosU_{isCAN}$ to form a stable homdimer in isolation is a considerable advantage, since it removes the presence of the homodimeric complex as a potential off-target. It is therefore only able to form a stable CC when combined with cJun.

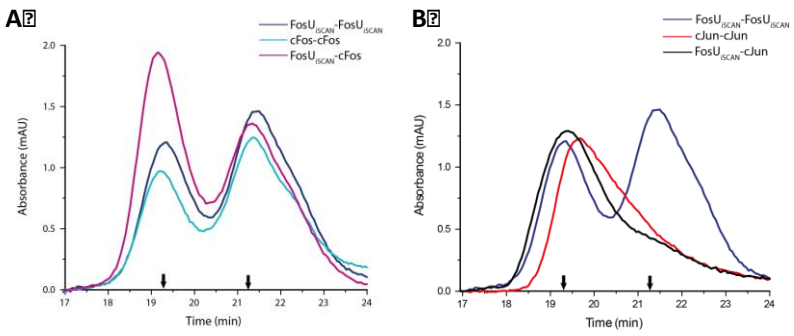

Figure 5 – Size Exclusion Chromatography Experiments. Shown are SEC profiles for post-melt samples. (A): Peaks at approximately 19.3 min and 21.3 min representing a mixture of dimer and monomer respectively, in the $FosU_{isCAN}$-cFos mixture (dark blue). Component cFos and $FosU_{isCAN}$ homodimer peaks show larger monomeric peaks compared to dimeric peaks. (B): $FosU_{isCAN}$ – cJun mixture generated a broad peak at approximately 19.3 with constituent cJun homodimer generating a peak at approximately 19.5 min, both indicating a dimer. Arrows show previously characterised controls on 39-mer peptides*.

The difference between the experimental thermal stability values and the values predicted by isCAN (Fig 3 vs Fig 6) is of interest. Observing the desired complex $FosU_{isCAN}$ – cJun, there was a decrease of 34°C between the predicted and experimental values. This is similar for 4hFosW and 4hFosW-cJun[24], where the complex was measured to be 16-17°C higher than bCIPA predicted. However, there was an observed decrease in the stability of the predicted extended cJun-cJun interaction which, combined with the off-target complexes, was not found to form a stable interaction. Overall, this means that the ΔT_m value has dropped reduced from 50°C predicted to 30°C when measured. Although a significant decrease, the measured value represents a bigger different between desired state stability and nearest off-target stability than that for any previous inhibitor peptide we have developed. Previous work exploring the further biophysical characteristics of peptides with similar thermal stability through isothermal calorimetry[24] gives insight into the importance of this difference. FosW-cJun (T_m = 63°C) displayed a K_d value of 39 nM whereas 20HC – cJun (T_m = 33°C) displayed a K_d of 15 µM. Since this has a similar thermal stability to the $FosU_{isCAN}$ – cJun interaction and the closest off target, cJun-cJun, we can estimate that the 30°C ΔT_m value denotes a sizable shift in the K_d from the range of nM (desired state) to µM (off-target states)[3,24].

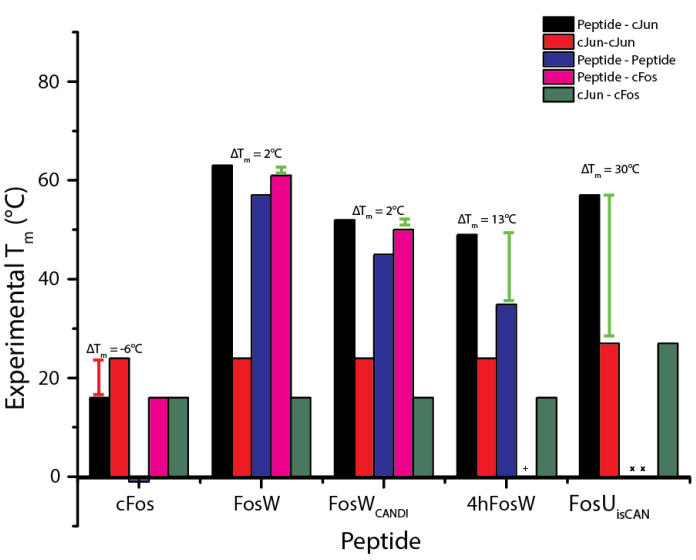

Figure 6. Measured T_m values of the isCAN selected peptide. A comparison of measured $FosU_{isCAN}$ with previously designed peptides FosW, $FosW_{CANDI}$ and 4hFosW. The T_m of the cJun-$FosU_{isCAN}$ was measured as 57°C with a ΔT_m of 30°C from the cJun-cJun homodimer Tm of 27°C. [x]Neither the $FosU_{ISCAN}$ nor mixture with cFos was found to form dimers, with the thermal denaturation profile unable provide a measured T_m (T_m < 0°C). [+]4hFosW-cFos thermal denaturation data is missing. (4hFosW homodimer data previously unpublished).

bCIPA Screening - The discrepancies between T_m values predicted by isCAN and the measured values validated through CD suggest that predictions are overestimating the stability of some complexes, which are generally higher than experimentally measured thermal melts. A simple reason for this is the scope of the underlying bCIPA and how it was developed. Using the interaction profiles of 45

peptides (and tested on 59^2 interactions), it means that bCIPA has a wide scope for predicting peptide interactions[3]. However, there are some instances in which the algorithm does not have the required data to estimate a contribution to binding affinity. For example, a Lys-Ala is not estimated to make a contribution. Compared to a known and quantified interaction, it is an example of a non-optimal interaction. In comparison to interactions that are known to be non-favourable (and thus positively scored), a value of 0 is seen as more favourable – although it represents a lack of data. This may explain some of the discrepancies in predicted/actual T_m values observed, with the incorporation of energetically non-favourable interactions. This is mitigated somewhat by the inclusion of the helical propensity values that each residue contributes – meaning that if the electrostatic contributions assumed within the algorithm are incorrect, there are other parameters that the algorithm uses to select residues.

From a software development perspective, the version used within this work stores all of the library sequences within the memory of the program, as well as data generated by interaction calculations. Further development is ongoing, with the goal of minimising the amount of data stored within the active program at any one point. It is hoped that this will limit the computational expense of this software and remove an obstacle in scaling up for high-performance systems. This would allow the isCAN approach to be used with larger systems and increased numbers of off-target peptides. Other *in silico* approaches with peptides have made use of other forms of searching within a large data set (including genetic algorithms[32] and Monte Carlo methods[33]). These methods of searching are typically coupled with molecular dynamics and docking simulations. If applied to a pairwise search with appropriate methodology, this could be a novel way to further screen for suitable peptides. Moreover, as we have previously demonstrated[6], bCIPA can be trained for specific bZIP subsets to increase its accuracy in such systems. Where knowledge of binding affinities is scarcer and more accuracy is required, an approach could be employed where exploratory experiments are conducted and the data used to create the necessary bespoke training/test sets. This could allow the predictors within the algorithm to be adjusted accordingly for the bZIP profile.

Limitations of Pairwise-Approach - It has been suggested that the interaction of residues might not be limited to the pairwise model that bCIPA uses[19]. Experiments that computationally derive additional scoring mechanisms from reported interaction affinities found that "triplet scoring" – the concept that the combinations of 3 residues at contact positions between **a**, **d**, **e** and **g** – could play a role in the prediction of coiled coil interaction.[19] Evidence that a combination of pairwise and triplet predictors increase accuracy provides further support for our "charge block" concept (blocks of same charge electrostatic residues at e and/or g postions)[6,7]. The charge block observations correlate to more nuanced, context specific stabilisation due to **g**/**e** residues interacting with **a**/**d**

residues to modify total interaction, the underlying algorithm could be improved to reflect this. As described above, previous work on peptides of a particular profile[6] has shown that training the algorithm on similar sequences has further optimised the weighting of the predictors to better predict T_m values. Since our studies have focused on Fos/Jun family bZIPs, a similar technique could be applied here. However, the lower T_m from predicted to measured values is consistent across the many $FosU_{ISCAN}$ interactions studied in this system. This suggests that, although non-optimal for our elongated peptides, the software is still able to correctly predict T_m relationships. In comparison with previous work in this field – peptides generated solely through PCA and CANDI[3,13] – this marks significant progress. As can be seen in Figure 6, although there is less than 10°C between the predicted T_m of cJun-4hFosW and cJun-$FosU_{isCAN}$, there is a measured increase of 17°C in the ΔT_m, making $FosU_{isCAN}$ much more specific than 4HFosW for cJun relative to Fos. This value – indicative of the ability to design against negative and competitive states – while maximising desired state stability shows the real strength of the isCAN technique. Our aim was to create a competitive binder for cJun which, through high-throughput computational screening, would address conflicting design requirements between desired and undesired states. The increased ΔT_m of $FosU_{isCAN}$ relative to previously designs, coupled with a high thermal stability with cJun, supports our initial hypothesis that *in silico* screening of peptides to mimic and control the parameters of a PCA-CANDI can result in peptides that can selectively inhibit cJun without interacting with cFos or other off-target bZIPs. The off-targets are then free to form transcriptionally active components of AP-1. Future exploration into combining this approach with an *in cellulo* PCA-CANDI would be the next step in validating and potentially generating useful antagonists for future peptide therapies targeting not only AP-1 dysregulation but any complex bZIP – mediated pathway in disease. This approach would provide a best-of-both combination of utilising very large libraries, screening via a computational approach to enrich for predicted binders with the desired attributes of high affinity and selectively, and then finally to experimentally screen the resulting reduced-size high quality library that is accessible to intracellular selection systems.

In conclusion, our work provides a framework by which bZIPs can be modelled within a CANDI environment with accuracy to derive highly selective peptide sequences. Driving the approach with a solely computational and data-driven framework allows us to collect data about peptide-peptide interactions and specificity both within and between bZIP families. As more and more experimental data becomes readily available, this approach will become increasingly valuable in the design of specific peptides that can target key components within increasingly complex bZIP interactomes.

AUTHOR CONTRIBUTIONS

J.M suggested and directed the research. A.L. developed the software, conducted the experiments, and synthesized, purified and characterised the peptides. Both authors participated in experimental design, analysis of the data, and writing the paper.

SUPPORTING INFORMATION

Information is provided on isCAN processing, tables of predicted melting temperatures (Table S1), experimental melting temperatures (Table S2), and peptide sequences. Software download information is also provided.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

ACKNOWLEDGEMENTS

JM is grateful to Cancer Research UK for a Career Establishment Award (A11738) and Pioneer Award (A26941), to the BBSRC (BB/R017956/1), and to the EPSRC for an Overseas Travel Grant (EP/M001873/2). JM and AL wish to thank the University of Bath for a PhD Studentship.

REFERENCES

(1) Mason, J. M., and Arndt, K. M. (2004) Coiled coil domains: Stability, specificity, and biological implications. *ChemBioChem 5*, 170–176.

(2) Fong, J., Keating, A., and Singh, M. (2004) Predicting specificity in bZIP coiled-coil protein interactions. *Genome Biol. 5*, R11.

(3) Mason, J. M., Schmitz, M. A., Muller, K. M., and Arndt, K. M. (2006) Semirational design of Jun-Fos coiled coils with increased affinity: Universal implications for leucine zipper prediction and design. *Proc. Natl. Acad. Sci. 103*, 8989–8994.

(4) Grigoryan, G., Reinke, A. W., and Keating, A. E. (2009) Design of protein-interaction specificity gives selective bZIP-binding peptides. *Nature 458*, 859–864.

(5) Grigoryan, G., and Keating, A. E. (2006) Structure-based prediction of bZIP partnering specificity. *J. Mol. Biol. 355*, 1125–1142.

(6) Crooks, R. O., Lathbridge, A., Panek, A. S., and Mason, J. M. (2017) Computational Prediction and Design for Creating Iteratively Larger Heterospecific Coiled Coil Sets. *Biochemistry 56*,

1573–1584.

(7) Crooks, R. O., Baxter, D., Panek, A. S., Lubben, A. T., and Mason, J. M. (2016) Deriving Heterospecific Self-Assembling Protein-Protein Interactions Using a Computational Interactome Screen. *J. Mol. Biol. 428*, 385–398.

(8) Park, W. M., Bedewy, M., Berggren, K. K., and Keating, A. E. (2017) Modular assembly of a protein nanotriangle using orthogonally interacting coiled coils. *Sci. Rep. 7*, 10577.

(9) Darnell, J. E. (2002) Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer 2*, 740–749.

(10) Rodríguez-Martínez, J. A., Reinke, A. W., Bhimsaria, D., Keating, A. E., and Ansari, A. Z. (2017) Combinatorial bZIP dimers display complex DNA-binding specificity landscapes. *Elife 6*, e19272.

(11) Crooks, R. O., Rao, T., and Mason, J. M. (2011) Truncation, randomization, and selection: Generation of a reduced length c-jun antagonist that retains high interaction stability. *J. Biol. Chem. 286*, 29470–29479.

(12) Baxter, D., Ullman, C. G., Frigotto, L., and Mason, J. M. (2017) Exploiting Overlapping Advantages of in Vitro and in Cellulo Selection Systems to Isolate a Novel High-Affinity cJun Antagonist. *ACS Chem. Biol. 12*, 2579–2588.

(13) Mason, J. M., Müller, K. M., and Arndt, K. M. (2007) Positive aspects of negative design: Simultaneous selection of specificity and interaction stability. *Biochemistry 46*, 4804–4814.

(14) Acerra, N., Kad, N. M., Cheruvara, H., and Mason, J. M. (2014) Intracellular selection of peptide inhibitors that target disulphide-bridged A??42 oligomers. *Protein Sci. 23*, 1262–1274.

(15) Cheruvara, H., Allen-Baume, V. L., Kad, N. M., and Mason, J. M. (2015) Intracellular screening of a peptide library to derive a potent peptide inhibitor of α-synuclein aggregation. *J. Biol. Chem. 290*, 7426–7435.

(16) Hagemann, U. B., Mason, J. M., Müller, K. M., and Arndt, K. M. (2008) Selectional and Mutational Scope of Peptides Sequestering the Jun-Fos Coiled-Coil Domain. *J. Mol. Biol. 381*, 73–88.

(17) Krylov, D., Barchi, J., and Vinson, C. (1998) Inter-helical interactions in the leucine zipper coiled coil dimer: pH and salt dependence of coupling energy between charged amino acids. *J. Mol. Biol. 279*, 959–972.

(18) Acharya, A., Rishi, V., and Vinson, C. (2006) Stability of 100 homo and heterotypic coiled-coil a-a' pairs for ten amino acids (A, L, I, V, N, K, S, T, E, and R). *Biochemistry 45*, 11324–11332.

(19) Potapov, V., Kaplan, J. B., and Keating, A. E. (2015) Data-Driven Prediction and Design of bZIP Coiled-Coil Interactions. *PLoS Comput. Biol. 11*, e1004046.

(20) Yang, Y., and Zhou, Y. (2008) Specific interactions for ab initio folding of protein terminal regions with secondary structures. *Proteins Struct. Funct. Genet. 72*, 793–803.

(21) Rohl, C. A., Strauss, C. E. M., Misura, K. M. S., and Baker, D. (2004) Protein Structure Prediction Using Rosetta. *Methods Enzymol. 383*, 66–93.

(22) Fields, G. B., and Noble, R. L. (1990) Solid phase peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acids. *Int. J. Pept. Protein Res. 35*, 161–214.

(23) Mason, J. M., Hagemann, U. B., and Arndt, K. M. (2007) Improved stability of the Jun-Fos activator protein-1 coiled coil motif: A stopped-flow circular dichroism kinetic analysis. *J. Biol.*

Chem. 282, 23015–23024.

(24) Baxter, D., Perry, S. R., Hill, T. A., Kok, W. M., Zaccai, N. R., Brady, R. L., Fairlie, D. P., and Mason, J. M. (2017) Downsizing Proto-oncogene cFos to Short Helix-Constrained Peptides That Bind Jun. *ACS Chem. Biol. 12*, 2051–2061.

(25) Fletcher, J. M., Bartlett, G. J., Boyle, A. L., Danon, J. J., Rush, L. E., Lupas, A. N., and Woolfson, D. N. (2017) N@ a and N@ d: Oligomer and Partner Specification by Asparagine in Coiled-Coil Interfaces. *ACS Chem. Biol. 12*, 528–538.

(26) Gurnon, D. G., Whitaker, J. A., and Oakley, M. G. (2003) Design and characterization of a homodimeric antiparallel coiled coil. *J. Am. Chem. Soc. 125*, 7518–7519.

(27) Oakley, M. G., and Kim, P. S. (1998) A buried polar interaction can direct the relative orientation of helices in a coiled coil. *Biochemistry 37*, 12603–12610.

(28) Thomas, F., Niitsu, A., Oregioni, A., Bartlett, G. J., and Woolfson, D. N. (2017) Conformational Dynamics of Asparagine at Coiled-Coil Interfaces. *Biochemistry 56*, 6544–6554.

(29) Monera, O. D., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993) Comparison of antiparallel and parallel two-stranded alpha-helical coiled-coils. Design, synthesis, and characterization. *J. Biol. Chem. 268*, 19218–27.

(30) Lau, S. Y., Taneja, a K., and Hodges, R. S. (1984) Synthesis of a model protein of defined secondary and quaternary structure. Effect of chain length on the stabilization and formation of two-stranded alpha-helical coiled-coils. *J. Biol. Chem. 259*, 13253–13261.

(31) ZHOU, N. E., KAY, C. M., and HODGES, R. S. (1992) Synthetic Model Proteins - Positional Effects of Interchain Hydrophobic Interactions on Stability of 2-Stranded Alpha-Helical Coiled-Coils. *J. Biol. Chem. 267*, 2664–2670.

(32) Heurich, M., Altintas, Z., and Tothill, I. (2013) Computational Design of Peptide Ligands for Ochratoxin A. *Toxins (Basel). 5*, 1202–1218.

(33) Russo, A., Scognamiglio, P. L., Hong Enriquez, R. P., Santambrogio, C., Grandori, R., Marasco, D., Giordano, A., Scoles, G., and Fortuna, S. (2015) In Silico Generation of Peptides by Replica Exchange Monte Carlo: Docking-Based Optimization of Maltose-Binding-Protein Ligands. *PLoS One* (Zhang, Y., Ed.) *10*, e0133571.